

Research Review Title: *Relationship Between Use of Quality Measures and Improved Outcomes in Serious Mental Illness*

Draft review available for public comment from July 10, 2014 to August 6, 2014.

Research Review Citation: Gaynes B, Brown C, Lux LJ, Sheitman B, Ashok M, Boland E, Morgan L, Swinson-Evans T, Whitener L, Viswanathan M. Relationship Between Use of Quality Measures and Improved Outcomes in Serious Mental Illness. Technical Brief No. 18. (Prepared by the RTI-UNC Evidence-based Practice Center under Contract No. 290-2012-00008-I.) AHRQ Publication No. 15-EHC003-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2015. www.effectivehealthcare.ahrq.gov/reports/final.cfm.

Comments to Research Review

The Effective Health Care (EHC) Program encourages the public to participate in the development of its research projects. Each research review is posted to the EHC Program Web site in draft form for public comment for a 4-week period. Comments can be submitted via the EHC Program Web site, mail or email. At the conclusion of the public comment period, authors use the commentators' submissions and comments to revise the draft research review.

Comments on draft reviews and the authors' responses to the comments are posted for public viewing on the EHC Program Web site approximately 3 months after the final research review is published. Comments are not edited for spelling, grammar, or other content errors. Each comment is listed with the name and affiliation of the commentator, if this information is provided. Commentators are not required to provide their names or affiliations in order to submit suggestions or comments.

The tables below include the responses by the authors of the review to each comment that was submitted for this draft review. The responses to comments in this disposition report are those of the authors, who are responsible for its contents, and do not necessarily represent the views of the Agency for Healthcare Research and Quality.

Commentator & Affiliation	Report Section	Comment	Response
Peer Reviewer #1	General	General Comments: I think this is a comprehensive and basically sound review. There are some papers relating quality measures and outcomes, e.g. Rosenheck RA, Fontana A, Stolar M (1999). Assessing Quality of Care: Administrative Indicators and Clinical Outcomes in Posttraumatic Stress Disorder. Medical Care 1999; 37(2)180-188. It's not clear why the review missed these. But I don't think they change the conclusion that no measures of SMI have been demonstrated to lead to better outcomes, although the proportion of people with schizophrenia receiving antipsychotic med is hard to argue with.	We appreciated the reviewer's assessment. The noted paper, while clearly important, addressed PTSD and did not meet our definition of SMI, so was not included.
Key Informant (KI) Reviewer #1	General	This is an excellent review of the literature on outcomes for individuals with SPMI, the utility (or lack thereof) of quality measures and the cost and feasibility of implementing QMs in the real world. And yet, as a clinician and health policy researcher, I am disappointed in this state of affairs. Today for people with SPMI we aspire for "recovery". We regret the bad outcomes of homelessness, incarceration, suicide, unemployment and economic dependency for SPMI individuals but the state of the quality measurement field is not up to the task of measuring recovery and incenting the treatment and the mental health system to achieve better results. The field is in a sorry state.	We agree with the reviewer that the area is understudied.
KI Reviewer #1	Additional Questions	Quality of the Report: Superior	We appreciate the reviewer's comment.
KI Reviewer #2	General	This is a useful review of quality measures for serious mental illness currently available. The study finds limited evidence linking these measures and health outcomes for patients with SMI, and recommends more research to better establish this link. Overall, the review is helpful but the question about whether these quality measures are linked with outcomes seems overly narrow. It would be helpful to review methods used by agencies developing these measures (e.g. NCQA, NQF) for assessing quality indicators (e.g. feasibility, relevance, scientific soundness) as well as simply the link with outcomes. It would also be important to assess whether quality measures in use in other medical conditions fare any better than those for people with SMI with regards to any of these criteria.	We agree with the reviewer that a critical review of methods to assess quality measures, but that charge goes beyond the scope of a Technical Brief. The reviewer's second point is also important. We have now added two citations in our "Summary and Implications" section, at the end of the third paragraph, that indicate that the state of the evidence in other medical conditions is not clearly better.

Commentator & Affiliation	Report Section	Comment	Response
Peer Reviewer #2	General	This report is extremely uninformative. it is laden with acronyms that are incomprehensible, it is poorly organized, and there is no general rationale for the whole review. The review does not address the major cost issue for severe mental illness, disability in everyday functioning. 75% of the total cost of mental illness is associated with functional deficits and disability it was actually amazing how little attention was paid to this.	The goal of this Technical Brief is to identify the current use of quality measures in the SMI population and to describe the evidence supporting their use. Cost is a key factor in the management of the seriously mentally ill, but the issue of cost was not in the scope of this Brief.
Peer Reviewer #3	General	I found the writing and organization of the document to be very clear. The questions, methods and results were appropriate and followed from the data. My main issue which made it difficult to read has to do with some of the logic. Is the implication that the absence of controlled research on the capacity for use of quality indicators to improve outcomes an argument for not measuring quality? There seemed to be a circularity to some of the logic. We have to use process measures because we can't measure outcomes; but use of these measures has not been shown to improve outcomes. This seems like a predetermined conclusion. What happens to the research on fidelity and outcome in EBP's? Also, perhaps this endeavor could look at what is known about measurement and implementation of EBP's. It seems that for this question, we may need to infer the importance of measuring quality from other types of studies. At the same time, the paper provided the foundation for this kind of speculation. The paper should use person first language throughout.	We appreciate the reviewer's thoughtful comments. The purpose of this Brief was to describe the current evidence base for the use of Quality Measures in the seriously mentally ill, and the results indicated that there was no evidence linking use of QMs with improved outcomes, nor evidence linking QMs with proxies for improved outcome. We do not believe that the absence of evidence argues for not measuring quality. We believe, as the reviewer notes, that these findings do provide the foundation to decide how one can move forward from the current state, and we identify key issues to help guide this movement.
Peer Reviewer #4	General	The aim of this Technical Brief is to identify how quality measures are currently used for SMI populations and the evidence supporting their use. The authors use a mixed methods approach that includes review of the literature and key informants. This is a very timely and important topic. The authors do a good job of describing the health care policy context and clinical importance of the topic. The brief is organized and the methods are clearly described. Their summary recommendations are thoughtful and i think will help move the field forward.	We appreciated the reviewer's comment.

Commentator & Affiliation	Report Section	Comment	Response
Peer Reviewer #4	General	<p>This brief can be improved upon in two categories: [#1 presented under Tables section]</p> <p>Context: The authors do not adequately provide context as to how some of the challenges in developing feasible, testable quality measures in behavioral health is shared more broadly in the quality measurement/development field. While I agree that there are some unique differences in measuring some quality domains for individuals who have severe mental illness compared to those with general medical conditions, there is also much that is shared. The report does not describe this at all, leaving the reader to think that there is more “exceptional” (and rudimentary) about quality measurement in behavioral healthcare vs general health care. Examples of how behavioral health quality measurement shares challenges with quality measurement in general are:</p> <p>Tension between measuring process measures and outcomes, and the risk that differences in outcomes merely reflect differences in patient case mix, rather than a direct impact of the care delivered .</p> <p>Process measures are typically derived by results from RCTs (which all share the same concerns regarding generalizability)</p> <p>In cases where face validity suggests an RCT would be unethical (eg, patients with disorder “x” should receive treatment “y” because the evidence is strong that “y” will lead to good outcomes), that development of process measures based on expert review of the literature are not only appropriate but also more ethical than conducting an RCT would be.</p> <p>The importance of feasibility and not overburdening providers.</p> <p>While I realize the focus of this brief is quality measures for severe mental illness populations, nonetheless, it would be better for the brief to at least mention that some of these challenges are not specific to quality measurement in this population, but are challenges for the field more broadly.</p>	<p>We agree and in the first paragraph of Summary and Implications note that the challenges of QM work are not limited to the SMI population. Points 2, 3, 4, and 5 are ones that we have already provided in the text for SMI populations.</p>

Commentator & Affiliation	Report Section	Comment	Response
Peer Reviewer #4	General	The acronyms GQs, QMs and QIs are distracting to the reader because they are specific to this brief and used frequently. I recommend they spell out these words.	We appreciate the reviewer's comments but will keep the acronyms as they are because we believe their use benefits more readers than it disadvantages. These acronyms are explained in first use, after which the acronyms are then kept in use.
KI Reviewer #3	General	This technical brief provides a useful, comprehensive review of the current state of practice on the use of quality measures and outcomes for persons with a serious mental illness. The brief addresses the major issues facing the field, and points in useful directions for further work aimed at establishing and incorporating such measures into routine practice. There are no glaring errors or omissions in the report.	We appreciate the reviewer's general assessment.
KI Reviewer #4	General	I have tried to organize my remarks by page and line number (with pagination referring to the number at the top of the page); however, the document includes frequent repetition of similar concepts. Many of the same remarks would also be applicable in other parts of the document and I have not been exhaustive in my cross referencing. I've included a number of examples to illustrate specific points -- I don't intend for these to be included in the report but hope they will help in understanding particularly complex concepts. These remarks are my personal opinions and should not be taken as representing the views of my employers.	. The brief has been substantially revised, and we believe our revisions address these comments. We note that there is some repetition, but we believe the repetition highlights key themes.
KI Reviewer #4	General	Overall, I think you've done an excellent job of discussing a complicated topic and reviewing the available information. I think it may be helpful to modify the title of the report, if possible. Quality measurement is really the focus but, with the current title, I would have assumed that the report related to approaches for measuring outcomes of serious mental illness or a summary of current evidence on prognostic outcomes of SMI. Thus, a title of "Relationship of quality measurement to outcomes of serious mental illness" (or something similar) that captures both of these concepts may be more informative to potential readers.	We appreciate the reviewer's comment. We have modified the title to read, "Relationship Between Use of Quality Measures and Improved Outcomes in Serious Mental Illness".
KI Reviewer #5	General	This was a well-written report covering a challenging topic. The questions were thorough but the report could be improved by placing the work within the context of recent discussions on the potential de-implementation of poor performance measures in chronic illness (see below).	We appreciate the reviewer's observations.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #6	General	1. Overall, I rated the report “fair.” Much of the writing came across as a collection of notes, rather than a coherent report. I found little in it that I felt I could use for my own work.	We acknowledge the reviewer’s comment. We identify what we think are critical issues for the field to address.
KI Reviewer #6	General	2. One problem was the failure to clearly define what was meant by “SMI.” This is a term that has been used for research and administrative purposes and does not have a single accepted meaning, variously being defined as particular groups of diagnoses, levels of service utilization, degree of disability, or some combination of these. Because of these, I was unable to fully determine the methods for identifying a measure as being one for SMI.	In our evidence review and in our Key Informant interviews we spend much time considering this definition. There are many definitions of SMI. As noted in our eligibility criteria, for the purposes of our Brief, we included populations “with SMI (define as a psychotic disorder, bipolar disorder, or MDD with psychotic features); or with serious and persistent mental illness (SPMI) but the specific psychiatric diagnose was not provided; or if SMI or SPMI is not specified but the patients had a psychiatric hospitalization.
KI Reviewer #6	General	3. Related to this, I thought it was a weakness that the paper did not discuss at all either the advantages/disadvantages of having measures specific to SMI (vs. those applying to all levels of mental disorder) or fully explain how a measure was identified as an SMI one.	The consideration of using a quality measure that is disease-specific vs. one that is more of general or global process measure is important, and this distinction organized our entire GQ 1b findings section. However, we did not discuss the advantages or disadvantages of specific vs. general SMI measures because neither that available evidence nor the KI interview identified this distinction as a key current issue.
KI Reviewer #6	General	4. For the questions concerning the context in which Qms are used, I was expecting to see information about their use by payers, such as state Medicaid and MH agencies, Medicare, and managed care organizations, and their incorporation into EHRs. There was little or nothing about this, which I attribute to what appears to have been an information collection methodology primarily focused on the academic literature.	Our literature search included both published peer-reviewed literature and gray literature (which includes government resources), and our key informants included multiple non-academic stakeholders. Neither our literature search nor literature identified by our Key Informants (which included multiple non-academic stakeholders) identified the information the reviewer mentions. Specifically, we looked for information regarding the uptake of QM use from both the literature and the KIs, and we did not find any.

Commentator & Affiliation	Report Section	Comment	Response
Janssen (Public Comment)	General	We welcome the opportunity to comment on this first of its kind document that provides a comprehensive evaluation of the state of quality measurement and the underlying scientific evidence supporting their use in the seriously mentally ill (SMI) population. We commend you for the development of this technical brief as it will aid health care decision makers in the application of quality measures into practice today and in setting a future measurement direction for this important population which faces high unmet need.	We appreciate the reviewer's comment.
Janssen (Public Comment)	General	First, we broadly support and agree with the full scope of content contained within the brief. Specifically, we emphasize our alignment with several key areas covered within the document. Need to explore, through research, the link between process measurement and patient outcomes in this patient population. Need to advance the range of psychiatric global, medical and patient-centered outcomes for development as quality measures.	We acknowledge the reviewer's support.
Janssen (Public Comment)	General	As you may be aware, NQF is exploring ways to elevate the nature of collaboration that occurs around new measure development through creation of a virtual "measure incubator." This kind of a forum would allow stakeholders interested in addressing measurement gaps in the SMI population to collaborate with measure developers, healthcare systems with an advanced state of measurement, quality collaboratives, and others. We suggest considering that the paper mention under the Next Steps section (i.e., 2. What are the most important outcomes with which to compare quality measures?), the potential benefit of engagement with a "measure incubator" forum, like the concept under consideration by NQF.	We highlight the need for this type of collaboration in our report and appreciate the NQF work in this area. We hope that this report may be helpful to them as they take the next steps.

Commentator & Affiliation	Report Section	Comment	Response
Janssen (Public Comment)	General	In the document, you briefly touched on recovery as a measurement and there may be interest in including additional depth of information on recovery upon awareness of such information. You may want to consider recent work by the Mental Health Center of Denver (MHCD). MHCD uses their consumer and provider-administered recovery measures as a measurement-based means to improve patient care and achieve recovery-based outcomes. Please note the recent publications of the scale validation papers by the Mental Health Centers of Denver for their consumer and provider-administered recovery measures. [1- Luszczakoski K, Olmos-Gallo PA, Milnor W, McKinney CJ, Measuring Mental Health Recovery, J Behav Health Services Res, 29May2014, 1-11; 2 – DeRoche K, Olmos Gallo PA, McKinney CJ, Starks R, Huff S, Measuring Recovery Related Outcomes: A Psychometric Investigation of the Recovery Markers Inventory, Comm Mental Health J, 18May 2014, 1-7.]	These are important publications. However, our aim was not to comprehensively review the validity of all key outcome measures, and these studies did not meet our inclusion criteria because they did not link QMs and outcomes.

Commentator & Affiliation	Report Section	Comment	Response
Janssen (Public Comment)	General	Section GQ 1e in the technical brief discusses the usefulness of currently available quality measures and calls attention to the importance of the development of an evidence database. We offer comment that the level of generalizability of quality measurement in order to meaningfully inform patient care is of critical importance and, thus, the use of pragmatic, real-world evidence, both observational (i.e., retrospective electronic medical records, registries) and interventional randomized controlled trials (i.e., pragmatic electronic medical records), in measuring/assessing quality measures for this population will be essential. Additionally, we offer consideration that the paper emphasize the need for the conduct of research on how quality measures affect clinician behavior and attitudes as well as administrator behavior and attitudes. We, therefore, overall recommend that the paper consider adding content to this section that speaks to the importance of real world evidence on the applicability of quality measurement on improved outcomes and overall patient care.	We agree that applicability to real world patients, clinics, and outcomes is essential, and our consideration of what outcomes are most important includes the need to have these outcomes matter to patients and providers in the real world. However, the GQ1e section reviews what literature is available to address the question of the usefulness of currently available QMs. Such a consideration is mentioned on page 20 GQ4b and in the Next Steps section. Also, in Next Step #3, we clarify the importance of real world evidence in our emphasis on the import of collecting data in real world settings: “the limited resource base to support implementation in real world clinic settings is a similarly important barrier. Indeed, the challenges of collecting QM data in ways that providers and organizations can use them; of analyzing the data; and of providing data-driven feedback to providers that allow a clinical response in real time may be as difficult as developing the evidence base. Key variables include personnel, electronic health records, and training on implementing and monitoring QMs.” Also, the report addresses how QMs affect clinician behavior in multiple places, including sections GQ1 b, c, and d and Summary and Implications (the third and fifth issues).
Janssen (Public Comment)	General	We thank you for your efforts to engage us and others as health care decision makers through this comment period in the development of this important technical brief to advance the state of quality measurement for the SMI population.	We appreciate the comment.
Peer Reviewer #1	Background	This all seemed logical and reasonable.	We thank the reviewer for the comment.
Peer Reviewer #2	Background	The review is particularly poorly written. The authors appear to adopt a strangely concrete approach to quality measures. It is largely incoherent and uninformative. The authors do not present a systematic description of how quality measures are developed.	The scope of our Technical Brief is the current state of evidence linking QMs and improved outcomes. Accordingly, we would not systematically include a review of how quality measures are developed.
Peer Reviewer #2	Background	There are a number of strange terms, such as “gray literature” that are discussed. The average reader and many expert readers such as myself will have no idea what the authors are talking about.	We have clarified what gray literature is (“i.e., non-peer reviewed material”) on page 4 in the Methods section, 1 st paragraph, 10 th line.
Peer Reviewer #3	Background	Good	We thank the reviewer for the comment.

Commentator & Affiliation	Report Section	Comment	Response
Peer Reviewer #4	Background	On page 1 (bottom), the authors state that quality measures general serve 3 purposes, one of them being research. That seems a bit confusing to me. I don't think the authors mean to state that one <i>purpose</i> of quality measures is research, but instead that there is research that uses quality measurement. Typically this kind of research is generated because of investigators having an interest in quality improvement and accountability. The way the language is currently regarding quality measurement and research seems to described an inversion of the relationship between the two.	We have clarified this point, modifying the text to read that QMs "are used for three general and often overlapping purposes" (which reflects better what the citation we reference notes). We have also removed the sentence sorting QMs into two broad categories, as this distinction seemed to confuse more than clarify.
KI Reviewer #1	Background	Excellent. But see general comments above.	We thank the reviewer for the comment.
KI Reviewer #2	Background	It would be helpful to provide more of a conceptual overview of quality indicators for people with SMI --some are measuring care for SMI and others are treating people with SMI as a high risk population at risk for medical problems. In the former case, how do measures address heterogeneity across mental conditions (e.g. schizophrenia vs. bipolar disorder vs. depression)? What are the tradeoffs between assessing conditions separately versus lumping them together? In the latter case, why were the particular medical measures chosen and how are they anchored in values for general medical populations (e.g. if rates of diabetes screening are low, are they worse than those seen in comparable populations without SMI)?	We have added a second paragraph under "The Use of Quality Measures in Serious Mental Illness" in the Background that address this point.
KI Reviewer #3	Background	Background information is well covered.	We thank the reviewer for the comment.
KI Reviewer #4	Background P. 11 lines 3-15.	I would suggest that the word "suicidality" not be used in these reports. I know that this term is often used as a global shortcut for suicide ideas, behaviors and death. However, this practice has contributed to much of the confusion in the literature on suicide risk assessment. (The word "suicidality" is commonly used in studies that only assess suicidal ideas or attempts and people infer that similar conclusions can be drawn about suicide, per se. However, the epidemiological characteristics and risk factors for suicide are very different from those for suicide attempts or suicide ideas.)	We have changed "development of suicidality" to read "development of suicidal ideas or behaviors".

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Background P. 11.	The discussion of patient centered outcomes is a bit confusing, especially the sentence at the end of line 9. Perhaps I've misunderstood, but I always thought that patient centered outcomes were intended to mean outcomes that were most meaningful to individual patients. Consequently, I assumed this would suggest considerable patient-topatient variability. Thus, symptom reduction might be patient centered for a patient with constant, belittling and threatening hallucinations or command hallucinations telling them to do something they viewed as horrible. On the other hand, symptom reduction might not be patient-centered for a patient with pleasant musical hallucinations. Patients may or may not value increased autonomy. (There is literature in non psychiatric populations showing that a sizable fraction of individuals do not wish active involvement in shared decision making; there is no reason to think this preference would be any different in psychiatric populations). Similarly, gainful employment may or may not be something that a particular individual desires.	We have added the points about meaningfulness and variability in patient outcomes in the Introduction.
KI Reviewer #4	Background	Maintaining employment is also mentioned as a clear demonstration of improved health -- I am not sure that is invariably the case. For example, some individuals can maintain employment despite severe symptoms as long as adequate supports are in place (e.g., formal supported employment programs, working in a family business with significant assistance provided, ongoing intensive treatment). Others can maintain employment but only at a much lower level of skill than their pre-illness baseline (e.g. a highly educated patient maintaining employment as a night shift mail sorter or restaurant dishwasher). Conversely, others are functioning quite well but cannot resume employment because they would lose their benefits and would not be able to maintain needed treatment (e.g. with clozapine).	We have added the points about meaningfulness and variability in patient outcomes in the Introduction.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Background	<p>Using functional metrics such as employment status for quality measurement can also be problematic because they can be altered by factors other than SMI. Individuals with SMI often have an illness onset before they finished their education or started a career. It is hard to determine what their functional status would have been in the absence of an SMI, particularly if other potentially impairing conditions are also present.</p> <p>For example, in the Suffolk County Psychosis project (Evelyn Bromet PI), we have followed a sizable cohort of individuals with SMI for 20 years. When we determine consensus ratings of standardized outcomes (e.g. functioning, employment status, GAF), it is extremely difficult to distinguish between illness related impairment and impairment due to premorbid characteristics. It can also be hard to distinguish between illness related effects and later unrelated events. As a hypothetical example, a man in his 40s is receiving disability for SMI, but had clear evidence of learning disability as a child and dropped out of school at age 16. He held a few odd jobs and then had a first break of a psychotic illness. In his early 20s, he was CONT'D IN NEXT ROW</p>	We agree on the challenges of measuring functional metrics and have previously noted this in the Introduction.
KI Reviewer #4	Background	involved in an auto accident and had an associated head injury with some apparent residual impairment. He has also smoked marijuana frequently for the past 20 years. This type of scenario is not unusual, but it becomes hard to parse out the relative contributions of each factor to ultimate poor work outcomes or other functional impairments. Without knowing these relative contributions, there is no way to know the extent to which the outcome relates to the patient's SMI or the extent to which it is modifiable.	
KI Reviewer #4	Background	Making providers accountable for quality measures that have multiple potential contributing factors and/or non-modifiable elements is problematic. Although this document does address some of the complexities of quality measurement relating to broadly defined outcomes, it may need greater emphasis and more specificity.	We agree, and this point is emphasized in GQ1b findings.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #5	Background	Overall, the background was thorough and it was appropriate to focus on research involving currently used measures. It would be helpful to place the work within the context of the Delphi panel methods used for the RAND QA tools study which included measures for depression, but similar issues arose in striking a balance between validity and feasibility (see McGlynn, Wells, etc.).	We have chosen to emphasize the current process for QM development, but these are important points: other studies, including work done by the RAND QA tool study using Delphi panel methods, have found similar challenges in striking a balance between validity and feasibility. We have now added this contextual information in the Background section.
Mariel Lifshitz (Public Comment)	Background [Page 12 of 124, line 23]	Does there have to be a single measure or metric? Is that a/the goal? Is having different metrics for different SMIs or settings necessarily bad/the consensus in the field? If not, I suggest revising to make that clear.	The goal is not a single measure or metric, but rather a list of potentially relevant QMs. This point is stated in the last paragraph of the Background.
Sunovion (Public Comment)	Background	We urge AHRQ to acknowledge and account for the role of prescription medicines in offsetting other areas of medical spending (such as hospitalizations and ER visits), as reported by the Congressional Budget Office (CBO) (among others), in addition to reducing SMI's costs to society that AHRQ has identified, such as lost productivity. Reference: CBO, "Offsetting Effects of Prescription Drug Use on Medicare's Spending for Medical Services" (Nov. 2012) (and other studies cited therein).	The emphasis in this brief was on the state of the evidence addressing whether use of QMs was related to improvement in outcomes in the SMI population. While cost is important, the focus of this Brief did not involve costs.
Sunovion (Public Comment)	Background	We support AHRQ's proposed focus on patient-centered outcomes and considerations. Evaluations of health care quality for patients with SMI often fail to focus sufficiently on patients' well-being and daily functioning; attention to these patient-centered aspects of care delivery, both from an outcomes perspective and a process perspective, is fundamentally important (p. 2).	We appreciate the reviewer's comment.
Sunovion (Public Comment)	Background	We appreciate AHRQ's recognition that certain Quality Measures (QMs) or uses of QMs could "lead providers to take on easier cases or penalize practices with more vulnerable populations" (p. 9; p. 22). To avoid this risk, we urge AHRQ to articulate clearly and explicitly that, because of the uncertainties about what is "best" in this area and the nascent state of the research, it is not appropriate to use QMs for SMI as a basis for coverage or reimbursement policies at this time.	We appreciate the comment. However, the purpose of the Brief was not to make recommendations or conclude about how to use QMs for SMI reimbursement, but rather was to review the state of the current evidence base and identify key issues.
Sunovion (Public Comment)	Background	We recognize the huge health care costs for managing SMI. Importantly, inpatient hospitalization for SMI has been recognized as one of the major cost contributors for healthcare costs. We recommend AHRQ to include literature of costs of inpatient hospitalizations among SMI (p.1).	The scope of this brief was on the state of the evidence addressing whether use of QMs was related to improvement in outcomes in the SMI population. While cost is important, the focus of this Brief did not involve costs.

Commentator & Affiliation	Report Section	Comment	Response
Sunovion (Public Comment)	Background	The cost of antipsychotic medications is one of the components of health care costs. We recommend AHRQ to assess the costs of antipsychotic medications from various perspectives. Apart from state pharmacy budget perspective, the use and costs of antipsychotic medications differ by patient types. Recent study (Desai et al. 2013) has shown that antipsychotic medications are the largest contributor (52%) among low-cost patients (mean annual costs of less than \$3,656) while inpatient hospitalizations are the largest contributor (28%) among high-cost patients (mean annual costs=\$27,944).	while costs is important, the scope of this brief was on the state of the evidence addressing whether use of QMs was related to improvement in outcomes in the SMI population. the focus of this Brief did not involve costs.
Sunovion (Public Comment)	Background	Substantial proportions of patients taking antipsychotic medications have comorbid conditions of “metabolic syndrome”. Patients consider weight gain issue as one of the most important and meaningful outcomes in their treatment decisions. We therefore recommend AHRQ to consider “metabolic syndrome” and weight gain are important meaningful outcomes (p.2).	These outcomes are important and are covered in the “Global outcomes” section of the findings in GQ1b.
Jacobsen SAMHSA (Public Comment)	Background	Title of the report needs to change to “Use of Quality Measures and Improved Outcomes in the Seriously Mentally Illness” Make sure the outcomes align with NBHQF (National Behavioral Health Quality Framework)	Thank you for the suggestion. We have modified the title to read, “Relationship Between Use of Quality Measures and Improved Outcomes in Serious Mental Illness”.
Peer Reviewer #1	Guiding Questions	These were quite comprehensive and I found them hard to follow, but generally I didn't disagree with anything that was said.	We acknowledge the reviewer's comment.
Peer Reviewer #2	Guiding Questions	Very poor definition of Quality measures. Complete inattention to literature on functional outcomes in schizophrenia	QMs have a variety of definitions. After reviewing the literature and discussing with our experts, we agreed on using this version. This definition was inclusive and considered the relevant literature on functional outcomes in schizophrenia.
Peer Reviewer #3	Guiding Questions	Good	We thank the reviewer for the comment.
KI Reviewer #1	Guiding Questions	Well done.	We thank the reviewer for the comment.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #2	Guiding Questions	Would make sure that these are specifically grounded in the conceptual framework and parallel those done in reviews of other conditions or populations.	No single well-established framework guides questions for SMI, but AHRQ has a typical template and framework for technical briefs, which we used for our GQs. As a result, our GQ are parallel to GQ in technical briefs of other conditions or populations. In addition, we have provided a comparison of our findings to the use of QMs in non-SMI populations in the Summary and Implications section (see above response to KI Reviewer #2 Druss in General section)
KI Reviewer #3	Guiding Questions	Well articulated, no comments.	We thank the reviewer for the comment.
KI Reviewer #5	Guiding Questions	The questions were extensive so it is difficult to determine whether additional questions should have been considered.	We acknowledge the reviewer's comment.
Sunovion (Public Comment)	Guiding Questions	Sunovion believes that AHRQ should consider if and how implementation of Quality Measures (QMs) for Psychiatric, Medical or Patient-Centered Outcomes should stratify findings by race/ethnicity, gender, age and other demographic variables, which could inform both individual- and population-level interventions to improve healthcare quality in the long-term (p. 10).	This point is an important consideration, but we found no available relevant literature.
Sunovion (Public Comment)	Guiding Questions	Sunovion believes AHRQ should evaluate inclusion of "patient reported outcomes" using "patient reported outcome metrics" as a separate category of outcomes or integrated within the existing three categories of outcomes, as an important complement to claims, survey and other types of data (p. 10). In the example list, Sunovion recommends AHRQ adding patient employment status, job placement, and return to pre-morbid occupational status (p.11).	The patient-related outcomes noted are important ones; however, how we list the outcomes in these pages reflects what the literature currently says and what the KIs reported.
Sunovion (Public Comment)	Guiding Questions	Three types of outcomes against which QMs are to be measured include: psychiatric, medical, and patient-centered outcomes (p.10). Since the QM definition was based on Donabedian's Structure-Process-Outcome model, it is logical and appropriate to include all three aspects of Donabedian's model against which QMs are to be measured. The current review focused only on the outcomes, but neglected the other two important component (e.g., structure of which healthcare delivery systems ensuring SMI care is accessible within a defined service area; processes of care delivery are appropriate). We recommend AHRQ include a review of the structure, process, and outcomes against which QMs are to be measured.	We carefully considered how best to apply this model to the use of QMs. We chose to focus on improved outcomes, rather than structure or process, because of the time and effort associated with wide dissemination, uptake, and maintenance of QMs. Given these resource considerations, we chose not to focus QMs that improve only process or structure (and not outcomes). Further, given what is available in the literature, we have also addressed the process dimension, as our intermediate outcomes (or proxies) monitor the process of care delivered.

Commentator & Affiliation	Report Section	Comment	Response
Sunovion (Public Comment)	Guiding Questions	Sunovion is concerned that a QM that discourages polypharmacy or focuses too heavily on medication count may discourage physicians from prescribing more than one psychotropic medication for fear of falling “outside recommended practices” (p. 12; see also p. 11). Since patients react differently to psychotropic medicines, or may experience evolving needs over time, physicians may need to adjust a patient’s regimen or introduce a new combination of medications. The American Psychiatric Association treatment guidelines for schizophrenia and for bipolar disorder both recognize the need for adjunctive therapy for many patients with SMI. (Lisa Dixon, et al., Guideline Watch (Sept. 2009): Practice Guideline for the Treatment of Patients with Schizophrenia, American Psychiatric Publishing, Inc, 2010; Robert Hirschfeld, Guideline Watch: Practice Guideline for the Treatment of Patients with Bipolar Disorder: APA Practice Guidelines, American Psychiatric Publishing, Inc, 2005.)	We note in the technical brief that QMs have both advantages and disadvantages. We do not endorse or discourage the use of any specific QM. Our goal is to review the available evidence linking QMs to outcome, not to develop QMs or to select which outcomes are proper/preferred or not.
Sunovion (Public Comment)	Guiding Questions	We applaud AHRQ’s proposal to include a medication adherence QM (p. 11). In considering adherence issues, it is important to note that many patients with SMI have poor medication adherence because they may lack insurance or, if insured, their plans often impose step-therapy or fail-first policies that require patients with SMI to first fail at older, less effective treatments before the insurer will cover a more effective medication. Many older medications have substantial side effects that can cause patients to miss doses or to stop taking their medication. We urge AHRQ to develop QM that would allow assessment of whether policies with step-therapy, fail-first, or other utilization management requirements result in lower medication adherence.	The purpose of this Brief was to describe the current evidence base for the use of Quality Measures in the seriously mentally ill and was not intended to develop or select QMs. Also, this list is not meant to be a comprehensive list of all QMs, rather it is meant to be a list identifying the evidence base for commonly used QMs.

Commentator & Affiliation	Report Section	Comment	Response
Sunovion (Public Comment)	Guiding Questions	While AHRQ includes prescription of a mood stabilizer as a specific proxy QM for patients with bipolar disorder, the proposed QMs do not track whether patients with bipolar disorder were prescribed an antipsychotic (p. 13). Antipsychotics are increasingly recognized as an important tool for management of bipolar disorder; tracking their use may help inform clinical decision-making or future development of treatment guidelines for bipolar disorder. (See, e.g., Robert Hirschfeld, Guideline Watch: Practice Guideline for the Treatment of Patients with Bipolar Disorder: APA Practice Guidelines, American Psychiatric Publishing, Inc, 2005.) We encourage AHRQ to add a QM addressing antipsychotic prescriptions for patients with bipolar disorder.	The purpose of this Brief was to describe the current evidence base for the use of Quality Measures in the seriously mentally ill and was not intended to develop or select QMs. Also, this list is not meant to be a comprehensive list of all QMs, rather it is meant to be a list identifying the evidence base for commonly used QMs.
Sunovion (Public Comment)	Guiding Questions	AHRQ has included patient voluntary discontinuation of medication as a QM for patients with bipolar disorder, but not for patients with schizophrenia (p. 14). We encourage AHRQ to add a schizophrenia-specific QM for patient reasons for voluntarily discontinuing medication.	The purpose of this Brief was to describe the current evidence base for the use of Quality Measures in the seriously mentally ill and was not intended to develop or select QMs. Also, this list is not meant to be a comprehensive list of all QMs, rather it is meant to be a list identifying the evidence base for commonly used QMs.
Sunovion (Public Comment)	Guiding Questions	GQ2: The report indicated that there were “no articles in the literature that predict the selection of QMs based on organizational components or characteristics of a hospital or other mental health care delivery (2a), ...or nonclinical patient characteristics (2d)” (p.17). One plausible explanation for non-existent literature on understanding contextual factors may be due to the inherit nature of this question. Contextual analyses of the use QMs require analyses across different settings. Existing QMs literature focuses reporting data within the same organization or database which preclude contextual analyses. To understand the contextual issues for which QMs are used, a synthesis of the literature may help address this data gap.	The report is a Technical Brief that summarizes the issues, not a Systematic Review that synthesizes the data. The decision as to whether to commission a full systematic review is outside the scope of our Technical Brief.
Sunovion (Public Comment)	Guiding Questions	We noticed some formatting issues on regarding citations 26, 32, 41(line 1 of GQ2e) and citations 26, 41 (line 5 of GQ2e). These citations need to be superscripted (p.17).	Thank you. We have corrected these errors.

Commentator & Affiliation	Report Section	Comment	Response
Sunovion (Public Comment)	Guiding Questions	We appreciate AHRQ's discussion of issues arising from comorbidities and co-occurring conditions (e.g., pp. 1, 20, 21), which can arise either independently or as side effects for certain patients from particular types of treatments. We encourage AHRQ to continue focusing on the non-interchangeable aspects of different SMI treatments from the perspective of tolerability and side effects issues for individual patients. These are important, patient-centered considerations and outcomes to account for when deciding on or evaluating a treatment.	We acknowledge the comment by the reviewer.
Sunovion (Public Comment)	Guiding Questions	Sunovion commends AHRQ's recognition of ongoing research assessing the impact of annual coverage limits on psychotherapy visits and other medical services, as well as the impact of changes in Medicaid psychiatrist fees (p. 21). These are important examples of the types of QMs that AHRQ should establish, as limits on coverage and payment for mental health care services can result in harmful under treatment for SMI.	Thank you for the comment.
Peer Reviewer #1	Methods	The authors reviewed an exhaustive set of sources. As noted above they missed some relevant papers that I am aware of, but I don't know why and didn't see obvious flaws.	We acknowledge the comment by the reviewer. We indicate the reason why any full text articles were excluded in Appendix C.
Peer Reviewer #2	Methods	Poor. The authors do not know what they are searching for and as a result do not find anything.	The Methods are consistent with best quality Technical Brief practices.
Peer Reviewer #3	Methods	Good	We thank the reviewer.
KI Reviewer #1	Methods	Complete as far as I can tell but I am not a methodologist.	We thank the reviewer.
KI Reviewer #2	Methods	Focus on grey literature is appropriate.	We thank the reviewer.
KI Reviewer #3	Methods	The methods for conducting the review were adequately described.	We thank the reviewer.
KI Reviewer #5	Methods	Overall, the methods were thorough. It might be helpful to rate the performance measures on the RAND criteria (validity, feasibility, clinical importance) perhaps using a Delphi panel if feasible.	This point is Important, but addressing it is beyond the scope of a Technical Brief because we did not use Delphi methods and did not rate the measures on RAND criteria..
Mariel Lifshitz (Public Comment)	Methods [Page 14 of 124, line 38]	Is there a start date for the lit search?	We have noted in our Methods Section (under Published Literature Search) that our literature search extends from inception through July 2014.

Commentator & Affiliation	Report Section	Comment	Response
Sunovion (Public Comment)	Methods	Sunovion appreciates AHRQ's inclusion of a patient advocate on the Key Informant (KI) panel. We are concerned, however, that the panel included only one patient advocate. We urge AHRQ to consult with more patient advocates in its efforts to develop patient-centered QMs, particularly given the importance of patient experiences and perspectives in achieving high-quality outcomes, as emphasized in the report (e.g., p. 15).	We agree that patient advocate input is key. We were pleased to be able to include one as one advocate given restrictions on the size of the Key Informant Panel.
Peer Reviewer #1	Findings	The authors apply very demanding standards and found the literature quite wanting. The requirement that use of measures be shown to improve outcomes is a very strict one, and not surprisingly measures for SMI have to meet this test. Have measures in other areas of medicine met this test? It would be useful to know a little more about successes elsewhere.	This point is important—and it is not clear that other areas of medicine have met this test. We make this point in our first paragraph of Summary and Implications. “We have attempted to describe the current state of the art regarding the use of QMs in the seriously mentally ill and to describe the evidence supporting their use. We note that the challenges of QM research are not limited to measurement in the SMI population, but extend to QM measurement in general.”
Peer Reviewer #2	Findings	There are no findings.	We were not able to identify any articles that met the inclusion criteria. Based on a review of related peer-reviewed articles, a grey lit search, and input from our KIs, we summarize our findings on result section.
Peer Reviewer #4	Findings	On page 15 of the report the authors state “Although commonly used, administrative data can be limiting owing to its lack of translational validity and also can have questionable validity because of unintended consequences of poor implementation.” It is unclear what the authors mean by lack of “translational” validity. Also, the statement about “questionable validity because of unintended consequences of poor implementation” would be a problem for all quality measures, not just those based on administrative data.	We agree the concept of translational validity is confusing, and we have replaced this phrase “the challenges of translating this form of data into clinically meaningful actions”. Also, we have removed the statement about “questionable validity” specifically being an issue with use of administrative data.
Peer Reviewer #4	Findings	The summary statement that no quality measure has been demonstrated to be associated with improved outcomes in this population is not correct. See Rost et al(11). Admittedly, this is one study and more are needed. Still, it's not correct to say “there are none out there.” [11: Rost K, Dickinson LM, Fortney J, et al.: Clinical Improvement Associated with Conformance to HEDIS-Based Depression Care. Mental Health Services Research 7:103-12, 2005]	In our population, those with SMI, we could not find a relationship between QM use and improved outcome. We did not include this study because it involves a population with depression, a number of whom have minor depression (hence not meeting our eligibility criterion requiring the population to have SMI).

Commentator & Affiliation	Report Section	Comment	Response
Peer Reviewer #4	Findings	Page 18: what do the authors mean by the “inherent unreliability in SMI diagnoses”? Are they referring to concerns about inter-rater reliability of diagnoses among clinicians? If so, at least for bipolar and schizophrenia, claims data have demonstrated to be accurate for establishing schizophrenia and bipolar disorder diagnostic cohorts, compared to structured clinical interview or chart review(12, 13). [12. Lurie N, Popkin M, Dysken M, et al.: Accuracy of diagnoses of schizophrenia in Medicaid claims. Hospital and Community Psychiatry 43:69-71, 1992 13. Unutzer J, Simon G, Pabiniak C, et al.: The use of administrative data to assess quality of care for bipolar disorder in a large staff model HMO. General Hospital Psychiatry 22:1-10, 2000]	Thank you for the suggestion. we have now changed the language to refer to “ <i>the potential for unreliability in SMI diagnoses</i> ”
Peer Reviewer #4	Findings	Page 21: typo—4th line from bottom of GQs 4d and e section: I believe the authors mean “Medicaid” rather than “Medical”.	We thank the reviewer and have made this correction.
KI Reviewer #1	Findings	Well stated	We thank the reviewer.
KI Reviewer #2	Findings	Good review of existing measures.	We thank the reviewer.
KI Reviewer #3	Findings	Generally, the findings appeared accurate. However, a number of corrections or considerations are suggested: 1. On page 9, concerning the use of claims data to evaluate receipt of evidence-based treatments, it should be noted that such data are limited by the fact that many evidence-based treatments do not have specific billing codes that distinguish them from non-empirically supported interventions.	We have now added this point in GQ 1a.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #3	Findings	2. On page 10, the distinction between psychiatric and patient-centered outcomes may not be as valid or correct as implied (e.g., symptoms vs. functioning). The authors do acknowledge that psychiatric outcomes and patient-centered outcomes may overlap, for example in the case of functioning, but then suggest that its measurement is of greater concern to patients and families than mental health treatment providers. Level of functioning must be assessed by treatment providers in order to match service intensity to patient need, and thus at a gross level providers are in fact invested in measuring functioning (often using global measures like the GAF). The distinction between psychiatric and patient-centered outcomes with respect to functioning may be in terms of the importance of specific different domains of functioning, such as work or self-care skills. Also, the sentence, “Mental health providers traditionally administer psychiatric outcomes...” is missing something. Administer assessments of psychiatric outcomes?	We have softened this implications by modifying the text to read that patient-centered outcomes “often are” of greater interest to families and patients. We have added the missing word (now reads “administer psychiatric outcome measures”).
KI Reviewer #3	Findings	3. On page 11, for global outcomes, why is “co-existing depression or anxiety symptoms” listed, but not the broader range of psychiatric symptoms, such as psychotic, negative, and manic/hypomanic symptoms? On the next page it is stated that psychotic symptoms (for schizophrenia) would captured by the global measures, as would manic symptoms (for bipolar disorder), but it isn't clear where.	The reviewer makes an important point. We reviewed our approved QM list, and identified the correct references that address the symptomatology for schizophrenia and bipolar monitoring, respectively.
KI Reviewer #3	Findings	4. On page 12, specific outcomes are considered for schizophrenia and bipolar disorder, but not major depression—why?	We have clarified this absence by stating in the text, “Of note, while there are a number of approved general measures for MDD symptoms, {CITE National Quality Measures Clearinghouse} we did not identify any approved QMs that measure depression symptoms in patients specifically having MDD with psychotic symptoms (meeting our definition of SMI).

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #3	Findings	5. Page 12, the specific proxy measures noted for schizophrenia do not appear to reflect a broader understanding of evidence-based psychosocial treatments for the disorder, such as family psychoeducation, assessment/treatment of co-occurring substance abuse, or supported employment. Assertive community treatment and case management referral should be included for bipolar disorder as well. In addition, the psychosocial interventions listed above for schizophrenia also have an evidence base supporting their effectiveness in bipolar disorder.	These measures are important. All are already listed under Psychiatric Global Proxy Measures, except for family psychoeducation (which is listed under Patient Centered Outcomes). Also, the reviewer is correct that referral for case management and for assertive community treatment are measures for bipolar disorder; we have now indicated this by place referral for each of these two strategies into Psychiatric Global Proxy measures.
KI Reviewer #3	Findings	6. Page 12, “examination of cognitive functioning and/or referral for cognitive behavioral therapy for psychosis” imply that they address the same problem, but they don’t. Impaired cognitive functioning would lead to recommendation for cognitive remediation (an empirically supported intervention) aimed at improving cognitive abilities in areas such as attention and concentration, memory, and executive function. Cognitive behavioral therapy for psychosis targets problematic symptoms (e.g., psychotic, negative) using a psychotherapy model that focuses on identification and modification of inaccurate and self-defeating beliefs or thinking patterns to more accurate thinking, and teaching more adaptive skills (e.g., interpersonal) for getting needs addressed. The reference for both statements (#42) appears to be a mistake—isn’t related to either cognitive behavioral therapy or cognitive remediation.	We have clarified that these two outcomes are different by separating them with a comma in our bulleted list. Also, we confirmed that the list reference (the National Quality Measures Clearinghouse) is correct.
KI Reviewer #3	Findings	7. On page 13, patient involvement in treatment should include both inpatient treatment AND outpatient treatment.	This is correct, but the QMs address outpatient treatment except for the one that is specified as inpatient. We have clarified this point in the text of GQ1b in the 6 th paragraph.
KI Reviewer #3	Findings	8. Page 13, receipt of help finding work is included as a specific outcome for schizophrenia, but should also be included for bipolar disorder (and major depression) as ample evidence supports the efficacy of this (i.e., supported employment).	We agree, we have moved this up to Patient Centered Global outcomes, since it covers both diagnoses.
KI Reviewer #3	Findings	9. Page 15, lack of training in assessment is another barrier to implementing quality measures among mental health providers.	We have added this point to the end of the first paragraph of GQ1d.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #3	Findings	10. On pages 15 or 16 it might be noted that billing codes that incentivize providers to give “treatment” more than “assessment” may be an obstacle to implementing quality measures.	While this point is likely true, we did not find this result in the literature, and since no other KI mentioned it, we will leave this point out.
KI Reviewer #4	Findings P. 18 lines 31-33.	It may be helpful to explain what this model is, since I don't think it's common knowledge. Also, only 2 literature citations are listed but the evidence is described as showing “consistent support”, which may overemphasize its importance.	References are available for readers to find out more about the Donabedian model if they choose to. Also, we have removed the word “consistent”.
KI Reviewer #4	Findings	P. 18 lines 33-41. An advantage of process based measures is that they tend to have fewer confounds (see discussion above) and do not have the same potential constraints on modifiability as outcome based measures. (Again, if providers are to be held accountable for outcomes, they need to have control over them.)	These points are already mentioned later in this section, so we will leave as is. Specifically, we note “As noted earlier, these outcomes—the ultimate goals for treatment—can be challenging to assess. Consequently, proxies for patient's health status (e.g., refills as a measure of adherence), that really evaluate <i>processes</i> of care or service use and are thought to be an <i>intermediate</i> step towards improved outcome, sometimes serve as outcomes against which QMs are evaluated. In the SMI population, more proxy (or process, or intermediate) measures are available, likely because they are more feasible and easier to collect. The likelihood of use varies somewhat by type of outcome.”

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Findings	To my knowledge there have not been any good studies of potential unintended consequences of using broader based outcomes as quality measures but I think these warrant consideration. If providers are held responsible for broad functional outcomes over which they have little or no control, they may avoid treating the most ill patients, the most disadvantaged patients or the patients who are least likely to adhere to treatment. Disparities of care are likely to increase and quality is likely to decrease rather than increase.	We discuss these points already in 1b and 1c. Specifically, we note in 1b, "Assessing these patient-centered outcomes represents a significant time and resource challenge, and may be especially difficult to improve in severe SMI groups, so using them to assess the effectiveness of QMs may be difficult and carry unintended consequences." In 1c, we note, "The literature also provided additional potential barriers of QM use in the SMI population, including (1) the redundancy of performance measurement efforts across several arenas; ^{41,50} (2) having questionable validity because of the unintended consequences of poor implementation (as noted above); ⁴¹ and (3) the limitations in using different types of data to characterize quality." We also note "Of note, despite KI concerns about how an emphasis on improved outcomes might lead providers to avoid the sickest SMI patients, the literature did not assess whether case-mix severity is a barrier to care for those with the most severe forms of SMI.
KI Reviewer #4	Findings	In terms of the link of performance measures to reimbursement, students clearly gravitate to fields with high levels of reimbursement. Tying mental health reimbursement to outcomes over which providers have limited control could make mental health specialties even less appealing for students than they already are. In the longrun, penalizing providers for such outcomes could also lead to broader workforce issues such as demoralization of providers, early retirements and further shortages of mental health professionals.	This point is plausible, but the purpose of this Brief was to describe the current evidence base for the use of Quality Measures in the seriously mentally ill, and we found no literature addressing it.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Findings	The use of suicide deaths as a quality outcome is not discussed in detail in the report but is a common focus of the Joint Commission and other regulatory agencies. Suicide death is obviously an important outcome with tragic consequences for patients, families and society. Clinicians can sometimes identify modifiable risks for suicide, with the presumption that intervention can reduce suicide risk. Unfortunately, other risk factors are nonmodifiable (e.g., age, sex, history of suicide attempt, history of psych admission) and other identified risk factors are not readily addressed by the health care system (e.g., housing problems, interpersonal difficulties, job loss). However, the approach of the Joint Commission and other regulatory bodies is that any suicide shortly after discharge is a sentinel event. This blames providers for an outcome that is statistically impossible to predict and that often has many non-modifiable contributors.	Suicide is certainly a sentinel and important event, but its rarity makes it hard to use as a quality measure, and we could not identify any relevant literature addressing this point.
KI Reviewer #4	Findings P. 19 line 25.	The experience with meaningful use (MU) among psychiatrists may be instructive here. Many of the MU measures are not metrics that make sense in routine psychiatric practice, at least with the frequency required by MU. Having to complete measures that are not part of the ordinary work flow can fragment care delivery and may lead to worse outcomes on other (unmeasured) factors. For example, the intent of handing out diagnosis specific education materials may be to increase patient engagement but for a patient in treatment for 10 years already, giving a simplistic handout about their diagnosis is not useful. Similarly, for a patient with no medication changes, giving new clinical summary at each visit takes time out of the visit to complete the medication reconciliation (to satisfy the meaningful use measure) and print out the summary. It also disrupts the flow of the visit in terms of psychotherapy. The quality measures that are available within meaningful use are similarly spotty in their coverage, don't always apply, and have other limitations. This is particularly true for psychiatric settings.	The purpose of this Brief was to describe the current evidence base for the use of Quality Measures in the seriously mentally ill, and we did not find literature that addressed meaningful use in this context.
KI Reviewer #4	Findings P. 19 line 48.	See discussion above. This might be one place to mention that outcomes such as level of functioning may need corrections for premorbid functioning or other factors (not just cross-sectional case mix adjustments).	We appreciate the suggestion, but we searched for this information and we did not find any relevant literature to address this point in the brief.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Findings P. 20 line 27-32.	This sentence may be easier to understand if split into two sentences. Also, I understand that the word “cases” is often used in clinical medicine, but it can also be viewed as impersonal or even dehumanizing. It may be preferable to avoid using the words “case” or “cases” when referring to an individual.	We appreciate the comment but believe it reads clearly as is. Also, we changed “cases” to “situations” here.
KI Reviewer #4	Findings Pp. 20-23	This seems like a very comprehensive listing. In a few situations, it wasn't clear to me why a measure was listed as outcome vs a proxy measure. For example, I would have thought that foot exams or eye exams in diabetics would be a proxy measure with rates of retinopathy or amputation due to foot ulcers being the corresponding global outcomes.	In this particular example, the exams themselves were not the outcomes, rather they define the relevant process of care, or proxy measure. While an outcome such as rates of retinopathy or amputation would be a meaningful outcome for this process measure, these were not reported in this population, so we could not list them. Rather, the outcomes listed in the literature for the diabetic population were HgA1C and LDL-C. In general, we provide a definition of these proxies on page 11 of the document (second paragraph of GQ1b): “proxies for patient's health status (e.g., refills as a measure of adherence), that really evaluate <i>processes</i> of care or service use and are thought to be an <i>intermediate</i> step towards improved outcome, sometimes serve as outcomes against which QMs are evaluated. In the SMI population, more proxy (or process, or intermediate) measures are available, likely because they are more feasible and easier to collect.”

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Findings	Under the patient centered outcomes, the issue of patient and family involvement in treatment is an important one but also serves as an example of the differences between a concept and its implementation. Thus, the Joint Commission seems to require an actual patient signature on the treatment plan despite the logistical difficulties of accomplishing this in many electronic records. By the same token, Joint Commission wants signatures of multiple treatment plan members of different disciplines on the treatment plan as well as descriptions of short and long terms goals, measurable objectives, etc. This complicated format is in contrast to more typical plans for treatment that are documented in progress notes in all other medical specialties. There are obviously many ways to involve patients in discussions about preferences, shared decision making and other aspects of treatment planning without using these time consuming processes. If treatment planning is framed as a quality measure, it would be important to have research on the optimal approach to use, particularly given the substantial amount of time that mental health professionals spend on such treatment plans already that could be better utilized in other modes of patient/family engagement or care. (I am not aware of any evidence that the style of treatment plan advocated by the Joint Commission and other regulatory agencies improves care or enhances patient/family engagement.)	This point is important. The scope of this Brief was to describe the current evidence base for the use of Quality Measures in the seriously mentally ill., not the specifics of which QM should be selected.
KI Reviewer #4	Findings P. 26 lines 26-42.	Footnotes should be changed to superscript font.	We have made this change. Thank you.
KI Reviewer #4	Findings P. 27 line 43.	This implies that QMs necessarily have to be linked to diagnosis, which I don't think is the case. One could have symptom focused QMs or treatment focused QMs (e.g. Weight and lipid monitoring in individuals on antipsychotic regardless of the indication). Other QMs may be relevant to anyone who is receiving treatment for a psychiatric indication (e.g. Assessment for substance use given high rates of comorbidity).	Here we are describing an issue on which there are competing arguments, and we have described the situation as mixed, so we will keep as is.
KI Reviewer #4	Findings P. 27 line 46-47.	See above discussion of patient centered outcomes. This statement implies that engagement is necessarily of most interest to stakeholders, which I'm not sure is the case.	This point reflects what KIs indicated in interviews. Accordingly, we have modified this to read "of most interest to many stakeholders"

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Findings	In terms of measuring recovery as a patient-centered quality outcome, there first should be a clear definition of recovery. However, in practice, there often seems to be confusion about this concept.	Thank you for the suggestion but we are not attempting to identify this as the most important outcome nor to provide operational definitions for the key outcomes.
KI Reviewer #4	Findings	I think it is important to delineate: 1. Remission from an episode of illness, with a return to one's normal state in the context of ongoing treatment (pharmacological and/or psychotherapeutic and/or other interventions) 2. Remission from an episode of illness with a return to one's normal state and maintaining of that normal state after treatment has ceased but with the understanding that ongoing monitoring is needed to detect potential recurrent episodes and intervene, if indicated 3. Full recovery in which the symptoms and signs of illness are gone and there is no anticipation of recurrence. This latter definition is how most patients would envision recovery even though it does not align with our knowledge of most psychiatric disorders and certainly does not align with most of our knowledge of serious mental illness.	Thank you for the suggestions. These are important points. The scope of this Brief was to describe the current evidence base for the use of Quality Measures in the seriously mentally ill.
KI Reviewer #4	Findings	These definitions have a number of implications for patients if health systems strive for this latter definition. For example, many patients feel they have somehow failed if they experience a recurrent episode and this is made worse if there is an ongoing focus of being in a "recovery-based" program. Health systems and state funded treatment or housing programs for those with SMI often have a vested interest in labeling individuals as "recovered" so that they can stop paying for treatment or supported housing. This can lead to a destabilization of patients' conditions, a need for acute treatment and a vicious cycle of "recovery" and relapse. A similar cycle can ensue if patients lack insight into the need for ongoing treatment to prevent relapse and emphasize their "recovery" as a reason to stop treatment.	Thank you for the suggestions. These are important points. The scope of this Brief was to describe the current evidence base for the use of Quality Measures in the seriously mentally ill.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Findings	With non-psychiatric disorders such as diabetes or hypertension, QMs do focus on maintaining a normal state (e.g., of glucose, BP) and minimizing complications of illness but there is no focus on “recovery” per se and no assumption that high quality care will result in a lack of need for further followup. With cancer, in which complete surgical removal could be said to result in a lack of signs/symptoms with no need for further treatment, metrics are expressed in terms of 5 year or 10 year disease free survival and not worded as recovery. Again, there is an unquestioned assumption that followup assessments will be needed to identify early signs of recurrence.	This need for follow-up in psychiatric and non-psychiatric disorders is important. We don't believe that our findings imply that high quality care will result in lack of need for further follow-up, nor do we believe that recovery as used by the psychiatric field implies no need for follow-up.
KI Reviewer #4	Findings	There is an implication in this section of the document that an outcome would be valuable as a QM if it is viewed as a desired outcome and is measurable. The missing link is that there also needs to be a demonstration that the quality of the delivered care actually influences the outcome. For example, care on an inpatient unit may be delivered in a high quality manner but if patient's degree of symptoms, level of imminent harm and lack of insight required involuntary admission, their perceived satisfaction and involvement in decision making is likely to be less than if they were admitted voluntarily. Individuals with high health literacy may have a greater ability to become engaged in their care and this could have a greater effect on engagement levels than the quality of care that was actually delivered.	We agree, and a main point of our findings is that there are a number of missing links—that between use of QMs and outcomes, that between use of QMs and proxies for outcomes. The missing link stated here, between proxies for outcomes and outcomes themselves, also may exist, but considering it is beyond the scope of this report.
KI Reviewer #4	Findings P. 28 line 51-53.	Other factors, especially those related to health care financing or regulations are probably more likely to affect QMs than quality improvement initiatives which usually have fairly small effects. Readmission penalties are one example and EMR meaningful use incentives are another.	In this section, we report on what the literature and our KIs described as key issues. The points noted here are important but were not identified by our literature search or our KI interviews.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Findings P. 29 line 34.	The fact that these are circumscribed data sets and that there are differences in the patient populations in these data sets (vs. all patients) also limits their generalizability. For example, patients who have Medicare related to disability are quite different in their illness severity than those with commercial insurance. Those over 65 yo with Medicare as a primary insurer are different from those with Medicare as secondary insurance. Those with Medicaid have a different socioeconomic profile, by definition, than those with Medicare or commercial insurance. Individuals with SMI may be uninsured even when they would otherwise qualify due to the complexities of the application process yet they would not be included in these administrative databases.	In this section, we report on what the literature and our KIs described as key issues. The points noted here are important but were not identified by our literature search or our KI interviews.
KI Reviewer #4	Findings P. 29 lines 31-46.	Is there a specific reason to believe that these interventions to improve cardiovascular or other health problems should be different in those with SMI vs. those without? Similarly, is there a reason to believe that preventive interventions shouldn't be at least comparable? (One might argue that more frequent screenings for cardiovascular or diabetic risk may be needed but not less.)	Possibly—in the Background, we note and reference that SMI populations die 10-25 years earlier than patients without these illnesses, primarily from cardiovascular disorders. Possible explanations may include both decreased access to care as well as decreased effectiveness of interventions (potentially related to other confounding factors). However, there is a lack of evidence addressing the effectiveness of these interventions to improve certain health conditions in the SMI population, and it is not clear whether this effectiveness differs from that found in non-SMI populations.
KI Reviewer #4	Findings P. 29 51-54.	It's not clear what is meant by "accountability" in this context. If this point is crucial, it may be useful to be more specific about what this framework means by the accountability.	We have defined "accountability" in the second to last paragraph of GQ4b, define it as "i.e., who is responsible for ensuring that QMs have an impact on improving care for those with comorbid medical problems."
KI Reviewer #4	Findings P. 30 lines 3-5.	Ethical prohibitions may also exist that would preclude some studies of QMs with high face validity. An example would be randomizing individuals to be asked or not asked about suicidal ideas.	Thank you for the suggestions. These are important points. , , This issue was not identified by literature search or on the KI interviews.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Findings P. 30 line 18.	Either in this item or a distinct item, it would be helpful to make specific mention of opportunity costs (i.e., what could be achieved by a different intervention or assessment using the same amount of time, personnel or other resources).	We have added this point in GQ 4c, where we note one area for future research to be “assessing the impact of QMs on the process of care and whether the benefits outweigh the costs (including opportunity costs)”. We did not further develop this idea because it did not come out in our literature review nor was it specifically mentioned by one of our KIs.
KI Reviewer #4	Findings P. 30 line 20.	There is an unstated assumption that QMs necessarily would be associated with greater cost efficiency. I don't think that assumption is necessarily correct. Also, cost efficiency for health plans (if they denied payment to providers) might be very different from cost efficiency for the health system as a whole or cost efficiency as viewed by patients, families or providers.	We have listed this area as an important one for future research, but we have not made any assumption about whether greater efficiency should be expected.
KI Reviewer #4	Findings	The statement about the EMRs seems rather non-specific, particularly compared to the level of specificity in the two prior items. At the very least, I would suggest changing this to read “determining the optimal ways to use data captured in EMRs to gauge the quality of care.” I don't think the use of EMRs for gauging quality is unique to health systems.	We have made this modification.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Findings	<p>You may want to split EMRs out as a distinct paragraph highlighting some or all of the following:</p> <ol style="list-style-type: none"> 1. Increasing use of EMRs poses opportunities for greater collection of structured data related to quality 2. Increasing impetus to capture structured data poses an increased likelihood of unintended consequences (e.g., focus on structured rather than narrative information leading to a loss of understanding of the patient as an individual, fragmentation of documentation leading to reduced ability to synthesize key information in clinical decision making, time required for structured documentation offsets benefits) 3. Problems with inter-rater reliability of structured data for subjective measures (which are typical in psychiatric practice and increasingly integrated into EMRs) 4. Challenges with patient rated scales being used for mandated quality measures due to factors that make scale completion difficult such as health literacy, language literacy, or psychiatric symptoms (e.g. Psychomotor retardation, catatonia, amotivation). 5. Potential for inappropriate conclusions to be drawn from EMR data if taken at face value without appreciating the source and confounds of the collated data (For example, if average hospital length of stay depends on the length of stay of patients discharged by a given provider, that provider could be the one penalized, even if they were more efficient than the prior provider who didn't discharge the patient.) 	<p>Thank you for the suggestions. These are important points. This issue was not identified by literature search or on the KI interviews.</p>
KI Reviewer #4	Findings	<p>Other EMR related considerations that might merit research: EMRs can also introduce systematic changes into a system that would affect quality measures (e.g., patient portals affecting engagement, provision of usual EMR educational materials may be positive for some patients whereas other patients may find them overly simplistic and unhelpful).</p>	<p>Thank you for the suggestions. These are important points. This issue was not identified by literature search or on the KI interviews.</p>

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Findings	Some aspects of EMRs and quality need to be specifically assessed for SMI patients. For example, it is not clear whether patient portal use and associated levels of patient engagement will differ for those with SMI as compared to portal use in individuals with other psychiatric disorders or non-psychiatric disorders. It is also not clear how EMR configurations intended to support privacy impact on the quality of integrated care or conversely how EMR features intended to support care integration affect patients' concerns about privacy or willingness to share information, which could impact the quality of delivered care.	Thank you for the suggestions. These are important points. This issue was not identified by literature search or on the KI interviews
KI Reviewer #4	Findings	The use of EMRs to issue quality related reminders is an additional quality issue that needs research to determine the best ways to deliver such reminders to providers to encourage adherence and avoid disruptions in workflow. EMR use by mental health professionals is likely to have unique workflow considerations, particularly when psychotherapy is being provided. The timing of quality related alerts may also need to differ compared to a typical medical visit. Qualitative research and time motion studies can be particularly useful in delineating optimal EMR workflows.	Thank you for the suggestions. These are important points. This issue was not identified by literature search or on the KI interviews
KI Reviewer #4	Findings P. 30 line 43.	The phrase that begins "changes in Medicaid..." seems to be missing a word.	We have corrected this wording.
KI Reviewer #4	Findings P. 32 line 3-4.	Such a shift in payment systems also has the potential for negative unintended consequences on care that need to be considered.	Thank you for the suggestion. These are important points. This issue was not identified by literature search or on the KI interviews
KI Reviewer #5	Findings	The findings were balanced but it was difficult to determine the next steps from the report. Moreover, the findings might be placed within the context of research by Sox-Harris (SUD measure validity and unintended consequences of performance measures). It would also be helpful the focus on de-implementing measures, which has been a recent issue in chronic illness (e.g., glucose, lipid control treat-to-target), and frame some of the responses to these questions around selection of valid measures (and de-implementation of invalid ones).	We believe the highest priority next steps are addressed in the Next Step section. The points here are important, but the key initial priority is to develop a sound evidence base.
Jacobsen SAMHSA (Public Comment)	Findings	I believe that this study is fair and appropriate in its findings.	We appreciate the feedback.

Commentator & Affiliation	Report Section	Comment	Response
Mariel Lifshitz (Public Comment)	Findings [Page 18 of 124, line 25]	Suggest not using “the serious mental illness population” if possible. We don’t want to lump everyone together. A suggestion would be to say “patients with SMI,” which is used on other pages in this TB, or something similar. This comment applies throughout this TB.	We have made this change in the specific spot mentioned. The other references to “serious mental illness” does not refer to “the serious mental illness population”.
Mariel Lifshitz (Public Comment)	Findings Page 23 of 124, line 29	Delete “of” between linking and QMs.	We have deleted this word.
Mariel Lifshitz (Public Comment)	Findings Page 24 of 124, line 7	Text reads “patent” but should be “patient”	We have corrected this part.
Mariel Lifshitz (Public Comment)	Findings Page 31 of 124, line 12	Delete “quality measure” before QM in the second sentence.	We have deleted this part.
Peer Reviewer #3	Next Steps	See above (General Comments)	
Peer Reviewer #4	Next Steps	<p>Recommendations: I agree overall with the recommendations in this report. However, the authors appear (perhaps) unaware of federal activities that have been underway in the past few years that directly address the issues identified in terms of improving the evidence base for effective interventions to improve medical outcomes for individuals with severe mental illness, as well as stakeholder forums to discuss quality measures. At a minimum, these efforts should be mentioned. For example: Recent NIMH RFAs for research aimed at testing interventions to improve medical outcomes for adults with severe mental illness and children with severe emotional disturbance. https://grants.nih.gov/grants/guide/rfa-files/RFA-MH-13-140.html http://grants.nih.gov/grants/guide/rfa-files/RFA-MH-14-060.html</p>	<p>These are important points, but not all directly are relevant to this report’s purpose.</p> <ol style="list-style-type: none"> 1. The RFAs noted, while addressing the SMI population, do not directly address quality measures. 2. The challenges of addressing the use of QMs in SMI is similar to challenges of addressing QMs in other medical illnesses; this is noted now in Summary and Implications. 3. We were aware of the Technical Expert panels mentioned and included a number of those participants are KIs and/or peer reviewers. 4. We appreciate the comments addressing the likely usefulness of non-RCT data, and we have added the potential role of quasi-experimental design studies to our first next step.

Commentator & Affiliation	Report Section	Comment	Response
Peer Reviewer #4	Next Steps	<p>Technical expert panels funded by ASPE and SAMHSA (of which I have participated in a few), that include various stakeholders (researchers, providers, health plans, state Medicaid mental health authorities, patients) to prioritize and define quality measures for individuals with severe mental illness. Several of these have been motivated by the ACA provisions requiring implementation of quality measures, the same provisions motivating this technical brief.</p> <p>Also, the proliferation of guidelines in behavioral health is no different than in general medical care. There are HEDIS, NQF, NICE recommendations in general medical care to name a few of the bodies that generate quality measures. If the authors feel that there is something different about behavioral health quality measures that requires one entity/organization agree upon a set of measures, it would be good for them to explain what that is/why it is so.</p> <p>Alternatively, if this is a challenge for the overall field of quality measurement (not specific to SMI quality measurement), it would be good for the authors to clarify that.</p> <p>I agree with the authors recommendation that quality measures require testing, and that for the field to advance, it will be important for studies that evaluate the link between a quality measure and outcomes. I also agree that one needs to be sensible and practical in that randomized controlled trials will not be the sole answer (and sometimes might be unethical for quality process measures that are solidly grounded in the evidence base as being associated with better outcomes), but that well done observational or quasi-experimental designs can be very informative as well.</p>	<p>These are important points, but not all directly are relevant to this report's purpose.</p> <ol style="list-style-type: none"> 1. The RFAs noted, while addressing the SMI population, do not directly address quality measures. 2. The challenges of addressing the use of QMs in SMI is similar to challenges of addressing QMs in other medical illnesses; this is noted now in Summary and Implications. 3. We were aware of the Technical Expert panels mentioned and included a number of those participants are KIs and/or peer reviewers. 4. We appreciate the comments addressing the likely usefulness of non-RCT data, and we have added the potential role of quasi-experimental design studies to our first next step.
KI Reviewer #1	Next Steps	Agree	We appreciate the reviewer's comment.
KI Reviewer #2	Next Steps	<p>Here is where it would be helpful to have a broader discussion of value of measures beyond discussion of whether they are associated with outcomes. Or is there something uniquely problematic for assessing quality for people with SMI? What are the tradeoffs between measures of "harder" outcomes such as hemoglobin A1C and "softer," but more patient-centered measures such as recovery? And to what extent are the limitations of these measures problems for all measures that are easily assessed using administrative data?</p>	<p>We appreciate this point. In Summary and Implications (1st paragraph), we have now indicated that the difficulties of addressing the evidence base for QM are not limited to the SMI population. We believe further discussion of the tradeoffs is beyond the purpose of this Technical Brief.</p>
KI Reviewer #3	Next Steps	Well thought through and articulated.	We thank the reviewer for the comment.

Commentator & Affiliation	Report Section	Comment	Response
KI Reviewer #4	Next Steps P. 32. Line 16.	Should be “mentally ill”	We have made this change.
KI Reviewer #4	Next Steps P. 32 line 21.	Suggest putting the initial phrase at the end of the sentence.	We have made this change.
KI Reviewer #4	Next Steps P. 32 line 23.	I think “practical considerations” is meant here.	This is correct, and we have made this change.
KI Reviewer #4	Next Steps P. 33 lines 20-22.	NY state has also used quality of care measure related to psychiatric care with their PSYCKES system. Molly Finnerty, who is a psychiatrist, has been actively involved in this work and could provide further details.	This may be true, but neither the literature review nor the KI interviews provided this information.
KI Reviewer #4	Next Steps P. 33 line 33.	Although there have been many quality measures proposed in various contexts for key governmental and regulatory QMs the options available for psychiatric clinicians are fairly sparse and tend to relate to a small number of circumstances and conditions (e.g., depression, antipsychotic use, ADHD). In the context of EMRs, vendors do not always have all allowable QMs available in their software. This leads to difficulties for psychiatrists’ compliance with programs such as meaningful use. Emphasizing the large number of available measures may be misleading, unless further qualifications to the statement are given.	This is a good point, and we have changed “available QMs” to “potentially relevant QMs”.
KI Reviewer #5	Next Steps	More discussion on the unintended consequences of performance measures, especially in light of the recent issues around the access to care measures in VA would be helpful.	This point about unintended consequences is important. We have pointed this out now in the first paragraph of Summary and Implications, where we discuss how these challenges are not limited to SMI populations.
Sunovion (Public Comment)	Next Steps	Sunovion commends AHRQ’s recommendation to include patients in developing effective QMs, as patient-centered outcomes should be the cornerstone of any QM development (pp. 23-24).	We thank the reviewer for the comment.
Sunovion (Public Comment)	Next Steps	AHRQ recommends “the design and implementation of a randomized controlled trial to test the effectiveness of a QM or set of QM relative to a meaningful and feasible outcome would be an especially important step.” While a randomized clinical controlled trial design is considered high-rigor with respect to balancing unmeasured factors, its usefulness in addressing the impact /effectiveness of QMs on outcomes is questionable. A prospective cohort study or disease registry study may present as more practical approaches to address this issue (p.23).	We agree, and this point is made in the “Suggestions” in Next Step # 1.

Commentator & Affiliation	Report Section	Comment	Response
Peer Reviewer #1	Summary and Implications	The one issue that the authors left out is cross-system use. SMI patients often get services from multiple providers so that no one health care system will have all the data needed to assess the care of its SMI patients. The VA comes closest to this but several studies have shown that even VA patients, especially those with co-morbid substance abuse use multiple healthcare systems, so readmission rates, or access to outpatient service measures are inevitably incomplete.	Cross-system use is an important point. We have now addressed this in our 4 th issue in Summary and Implications.
Peer Reviewer #2	Summary and Implications	Again, this review is so poorly conceptualized that there is no valuable information here.	We note the reviewer's comment.
Peer Reviewer #3	Summary and Implications	See above (General Comments)	
KI Reviewer #1	Summary and Implications	Nicely done.	We thank the reviewer for the comment.
KI Reviewer #2	Summary and Implications	Good summary of findings.	We thank the reviewer for the comment.
KI Reviewer #3	Summary and Implications	These follow well from the review--no suggested modifications.	We thank the reviewer for the comment.
KI Reviewer #5	Summary and Implications	The summary was comprehensive- can there be some additional discussion on the use of measures on de-implementation of non-evidence-based practices (see articles on the Choosing Wisely initiatives).	The point here is important, but the key initial priority is to develop a sound evidence base.
Sunovion (Public Comment)	Summary	We appreciate AHRQ's recognition that the value and appropriateness of QMs for SMI is important to explore further but depends heavily on context and on the purpose for which the QMs are being used. We agree that there is a need to address the current "absence of an evidence base" in this area and to consider, carefully and thoughtfully, possible options, methods, and research to advance the goal of potentially using QMs to improve the quality of care for individuals with SMI (p. 22).	We thank the reviewer for the comment.
Jacobsen SAMHSA (Public Comment)	Summary and Implications	What is missing seems to be an assessment of what will be required for meaningful change to occur.	We identify what we believe is required for change in the Next Steps that follows.

Commentator & Affiliation	Report Section	Comment	Response
Sunovion (Public Comment)	References	Please consider adding these citations onto the report. CBO, "Offsetting Effects of Prescription Drug Use on Medicare's Spending for Medical Services" (Nov. 2012). Desai P et al. J Manage Care Pharm 2013; 19(6): 468-77. Dixon L, Perkins D, & Calmes C. Guideline Watch (Sept. 2009): Practice Guideline for the Treatment of Patients with Schizophrenia. American Psychiatric Publishing, Inc, 2010. Robert Hirschfeld. Guideline Watch: Practice Guideline for the Treatment of Patients with Bipolar Disorder: APA Practice Guidelines, American Psychiatric Publishing, Inc, 2005.	We reviewed these against our eligibility criteria and none met selection criteria.
Peer Reviewer #4	Tables	This brief can be improved upon in two categories:[#2 presented under General section] 1. Literature review: the appendix tables describing the studies that use quality measures appears incomplete. It is not clear if the aim of the tables is to: 1) provide an exhaustive list of studies that use quality measures; 2) provide an exhaustive list of the quality measures used in research; or 3) provide an exhaustive list of quality measures used in research AND the research includes an intervention (so one could see if there was an improvement in quality based on the intervention. I recommend the authors clarify this. Regardless, none of these 3 goals (if any of those are the goals), is met. There is literature that they have missed (relevant to 1 and 2 above), and not all of their studies cited are relevant for #3 above. I don't think it's realistic to expect this report to have a complete compendium of all of the research that has used quality measures for individuals with SMI (#1 above). It's also not necessary. That said, if the authors would like to be more comprehensive and attempt to include all studies that incorporate quality measures (even if the measures are duplicative of those used in already cited studies), some examples are studies by Steve Soumerai, Stacie Duzetzina, Marguerite Burns, Vicki Fung, and Alisa Busch. This is not an exhaustive list. Just naming a few investigators this this area.	The purpose of these appendices is not to be comprehensive of all the literature addressing these outcome measures; rather, it is to be comprehensive of the outcomes measured. We have reviewed each of these articles, and we have included those that meet our inclusion criteria and add an important citation or an important outcome to our GQ1b outcomes. Accordingly, we have added the following 7 citations to the GQ 1b text and the GQ 1b table: 1) Druss BG, Zhao L, Cummings JR, et al. (added a citation to Global Medical outcomes) 2) Moeller KE, Rigler SK, Mayorga A, et al. (added a new citation to Global proxy Medical outcomes) 3) Valenstein M, Copeland LA, Blow FC, et al. (added a new specific Psychiatric proxy outcome to Schizophrenia group) 4) Valenstein M, Ganoczy D, McCarthy JF, et al. (added a new specific Psychiatric proxy outcome to Schizophrenia group) 5) Lehman AF, Steinwachs DM, the Co-Investigators of the PORT Project. (added a new specific Psychiatric proxy outcome to Schizophrenia group) 6) Busch AB, Frank RG, Sachs G: (added a new specific Psychiatric proxy outcome to Bipolar group) We have not added the following: 1. McGinty EE, Blasco-Colmenares E, Zhang Y, et al. (no specific use as a QM was mentioned).

Commentator & Affiliation	Report Section	Comment	Response
Peer Reviewer #4	Tables	<p>In terms of ensuring that the tables are complete regarding specific quality measures (which I think is consistent with their overall goals for the technical brief) I recommend the authors look to the literature on general medical quality monitoring for individuals with severe mental illness (e.g., work done by Benjamin Druss et al (1), Moeller et al (2), or Gail Daumit et al (3, 4)). Also, the literature on Medication Possession Ratios (5-8); in some studies MPR < 80% has been associated with hospitalization for individuals with schizophrenia. The seminal study on quality of care for individuals with schizophrenia, the Schizophrenia PORT, was not included in the review and should be (9). Another example of a quality measure used in research but not included in the tables was "no antidepressant prescribed in absence of a mood stabilizer" for individuals with bipolar-I disorder (10).</p> <p>[1. Druss BG, Zhao L, Cummings JR, et al.: Mental comorbidity and quality of diabetes care under Medicaid. Medical Care 50:428-33, 2012 2. Moeller KE, Rigler SK, Mayorga A, et al.: Quality of monitoring metabolic effects associated with second generation antipsychotics in patients with schizophrenia on public insurance. Schizophrenia Research 126:117-23, 2011 3. Cullen BA, McGinty EE, Zhang Y, et al.: Guideline-concordant antipsychotic use and mortality in schizophrenia. Schizophrenia Bulletin 39:1159-68, 2013 4. McGinty EE, Blasco-Colmenares E, Zhang Y, et al.: Post-myocardial-infarction quality of care among disabled Medicaid beneficiaries with and without serious mental illness. General Hospital Psychiatry 34:493-9, 2012 5. Sajatovic M, Valenstein M, Blow F, et al.: Treatment adherence with lithium and anticonvulsant medications among patients with bipolar disorder. Psychiatr Serv 58:855-63, 2007 6. Sajatovic M, Valenstein M, Blow FC, et al.: Treatment adherence with antipsychotic medications in bipolar disorder. Bipolar Disorders 8:232-41, 2006 7. Valenstein M, Copeland LA, Blow FC, et al.: Pharmacy data identify poorly adherent patients with schizophrenia at increased risk for admission. Medical Care 40:630-9, 2002</p>	<p>The purpose of these appendices is not to be comprehensive of all the literature addressing these outcome measures; rather, it is to be comprehensive of the outcomes measured. We have reviewed each of these articles, and we have included those that meet our inclusion criteria and add an important citation or an important outcome to our GQ1b outcomes. Accordingly, we have added the following 7 citations to the GQ 1b text and the GQ 1b table:</p> <p>1) Druss BG, Zhao L, Cummings JR, et al. (added a citation to Global Medical outcomes) 2) Moeller KE, Rigler SK, Mayorga A, et al. (added a new citation to Global proxy Medical outcomes) 3) Valenstein M, Copeland LA, Blow FC, et al. (added a new specific Psychiatric proxy outcome to Schizophrenia group) 4) Valenstein M, Ganoczy D, McCarthy JF, et al. (added a new specific Psychiatric proxy outcome to Schizophrenia group) 5) Lehman AF, Steinwachs DM, the Co-Investigators of the PORT Project. (added a new specific Psychiatric proxy outcome to Schizophrenia group) 6) Busch AB, Frank RG, Sachs G: (added a new specific Psychiatric proxy outcome to Bipolar group) We have not added the following: 1. McGinty EE, Blasco-Colmenares E, Zhang Y, et al. (no specific use as a QM was mentioned).</p>

Commentator & Affiliation	Report Section	Comment	Response
Peer Reviewer #4	Tables	<p>8. Valenstein M, Ganoczy D, McCarthy JF, et al.: Antipsychotic adherence over time among patients receiving treatment for schizophrenia: A retrospective review. J Clin Psychiatry 67:1452-550, 2006</p> <p>9. Lehman AF, Steinwachs DM, the Co-Investigators of the PORT Project: Patterns of usual care for schizophrenia: Initial results from the schizophrenia Patient Outcomes Research Team (PORT) client survey. Schizophrenia Bulletin 24:11-9, 1998</p> <p>10. Busch AB, Frank RG, Sachs G: Bipolar-I depression outpatient treatment quality and costs in usual care practice Psychopharmacology Bulletin 41:24-39, 2008]</p>	<p>The purpose of these appendices is not to be comprehensive of all the literature addressing these outcome measures; rather, it is to be comprehensive of the outcomes measured. We have reviewed each of these articles, and we have included those that meet our inclusion criteria and add an important citation or an important outcome to our GQ1b outcomes. Accordingly, we have added the following 7 citations to the GQ 1b text and the GQ 1b table:</p> <ol style="list-style-type: none"> 1) Druss BG, Zhao L, Cummings JR, et al. (added a citation to Global Medical outcomes) 2) Moeller KE, Rigler SK, Mayorga A, et al. (added a new citation to Global proxy Medical outcomes) 3) Valenstein M, Copeland LA, Blow FC, et al. (added a new specific Psychiatric proxy outcome to Schizophrenia group) 4) Valenstein M, Ganoczy D, McCarthy JF, et al. (added a new specific Psychiatric proxy outcome to Schizophrenia group) 5) Lehman AF, Steinwachs DM, the Co-Investigators of the PORT Project. (added a new specific Psychiatric proxy outcome to Schizophrenia group) 6) Busch AB, Frank RG, Sachs G: (added a new specific Psychiatric proxy outcome to Bipolar group) <p>We have not added the following:</p> <ol style="list-style-type: none"> 1. McGinty EE, Blasco-Colmenares E, Zhang Y, et al. (no specific use as a QM was mentioned)

Commentator & Affiliation	Report Section	Comment	Response
Counts, Mental Health America (Public Comment)	Tables	Mental Health America would like to support the further research and evaluative use of recovery-oriented outcome measures for SMI quality of care measurement. Measures included in the draft, such as autonomy, social support, and quality of life hold exciting possibilities for the future of quality management. Mental Health America would also like to raise community inclusion as another important patient-centered quality measure and direct the Agency to the following additional resources: Community inclusion measures: Kaplan K, Salzer M., & Eugene Brusilovskiy. (2012). Community Participation as a Predictor of Recovery-Oriented Outcomes Among Emerging and Mature Adults with Mental Illnesses: Psychiatric Rehabilitation Journal. 35(3), 219-229. Townley G. & Kloos B. (2009). Development of a measure of sense of community for individuals with serious mental illness residing in community settings: J Community Psychol. 37(3), 362–380. Townley G., Kloos B., & Wright P. (2009). Understanding the experience of place: Expanding methods to conceptualize and measure community integration of persons with serious mental illness: Health Place. 15(2), 520–531. Reviews of recovery-oriented measures generally: Khanan D., McDonald K. & Williams Neils C. (2013) Measuring Recovery: A Toolkit for Mental Health Providers in New York City: NYC Department of Health and Mental Hygiene. Campbell-Orde R., Chamberlin J, Carpenter J, (2005). Measuring the Promise: A Compendium of Recovery Measures, Volume II. Cambridge, MA: Human Services Research Institute.	This is an important point. We reviewed each article for eligibility per our selection criteria. We have now added Community Inclusion as a Patient-Center Global Outcome in the text, and have cited in the text and our Tables the following relevant and eligible article: Kaplan et al.
Mariei Lifshitz (Public Comment)	Tables Page 60 and 61 of 124, table title	Delete parenthetical at end of sentence. Does not appear to belong	We have corrected for these tables, as well as for Table D-3.
Peer Reviewer #1	Clarity and Usability	It seemed well structured, but was written in a rather obsessive and dull style. Unfortunately I don't have any recommendations for improving the writing. While I agree with the conclusions, I think the authors should note that not including monitors of SMI care is, at face value, a terribly unfortunate conclusion because it means that care of some of the most disabled vulnerable Americans serviced by the ACA will not be accountable. I think this is largely a function of how fragmented that care for people with SMI is in the US and elsewhere.	We acknowledge the reviewer's comment and appreciate the agreement with our conclusions. We have addressed the point about the importance of monitoring the quality of SMI care at the start of Summary and Implications.

Commentator & Affiliation	Report Section	Comment	Response
Peer Reviewer #2	Clarity and Usability	There is no useful information in this report. It is incoherent.	We acknowledge the reviewer's comments.
Peer Reviewer #3	Clarity and Usability	Good	We thank the reviewer for the comment.
KI Reviewer #1	Clarity and Usability	Very clear....a real challenge to the field.	We thank the reviewer for the comment.
KI Reviewer #2	Clarity and Usability	Writing style is clear and well-organized.	We thank the reviewer for the comment.
KI Reviewer #3	Clarity and Usability	The report is well structured and organized, and will be useful in informing future research in this area.	We thank the reviewer for the comment.
KI Reviewer #5	Clarity and Usability	Very clearly presented and well-organized	We thank the reviewer for the comment.
KI Reviewer #4	Structured Abstract –	P. 8 line 44. The current wording could imply that no studies had been done or that studies were available but showed no evidence for an effect. I would suggest being more specific in the abstract about what is meant here.	We have made this correction. We state now that “No studies have assessed whether the use fo QMs improves health outcomes . . .”.
Mariel Lifshitz (Public Comment)	Structured Abstract Page 8 of 124, line 14	QM was already defined in the first sentence, so don't have to redefine it here.	We have made this correction.
Mariel Lifshitz (Public Comment)	Structured Abstract Page 8 of 124, line 44	Need to make the connection between quality of care and quality measures earlier. Suggest inserting something in the background section that makes this connection. For example, can use the language from the first sentence of the fifth paragraph on page 1 or the first sentence of the third paragraph on page 2.	We have made this addition (3 rd sentences of Background).
Mariel Lifshitz (Public Comment)	Structured Abstract Pg. 8 of 124, line 9	insert “of” between use and validated.	We have made this correction.